

## 跨语言标注情境中的学习行为

——基于推荐标签的实验分析

Study on the Learning Behavior in the Context of Cross-language Tagging: An Experimental Analysis Based on Recommended Tags

张思雅<sup>1</sup> 马晓悦<sup>2</sup> 赵捧未<sup>1</sup>

(1. 西安电子科技大学, 西安, 710071; 2. 西安交通大学, 西安, 710049)

[摘要] [目的/意义]厘清跨语言环境中推荐标签及其表征语言对用户标注学习行为的影响,是优化和提高推荐标签质量、改善用户跨语言标注体验的重要基础。[研究设计/方法]通过受控实验的方法对比用户在外文推荐标签、中文推荐标签以及无推荐标签的情况下的跨语言标注结果和学习行为。[结论/发现]跨语言标注环境下存在三种基于推荐标签的学习模式:基于标签形式的学习、基于标签语义的长语义映射学习和基于标签语义的短语义映射学习;推荐标签在为用户的标注学习提供指导的同时,也带来一定的代价,这符合“锚定效应”的观点。[创新/价值]研究结果将有助于指导跨语言标注系统的设计与实现,提高信息共享效率。

[关键词] 跨语言 推荐标签 学习行为 信息标注

[中图分类号] G250 [文献标识码] A [文章编号] 1003-2797(2020)04-0076-11 DOI: 10.13366/j.dik.2020.04.076

[Abstract] [Purpose/Significance] Clarifying the effects of recommended tags and their representational languages on learning behavior in the cross-language environment is an important foundation for optimizing and improving the qualities of tag suggestion and enhancing user experience in cross-language tagging. [Design/Methodology] With the use of controlled experiment, this study compared the results and learning behaviors of cross-language tagging while users annotated texts in three modes, which are tagging with recommended tags in foreign languages, tagging with recommended tags in Chinese and tagging without recommended tags. [Findings/Conclusion] a. There are three learning patterns based on recommended tags in the context of cross-language tagging, including learning based on the tag forms, the long semantic mapping-learning based on the semantics of tags and the short semantic mapping-learning based on the semantics of tags; b. Recommended tags not only provide guidance for users in their learning of tagging, but also bring them some cost. It is consistent with the “anchoring effect”. [Originality/Value] The findings will be helpful for the design and implementation of cross-language tagging system, and also improve the efficiency of information sharing.

[Keywords] Cross-language; Recommended tags; Learning behavior; Information tagging

[基金项目] 本文系教育部人文社会科学研究规划基金“信息协同视角下基于可视化媒介的智慧应急响应交互行为研究”(19YJA870009)、陕西省自然科学基金基础研究计划-面上项目“复合视角下可视化标注系统驱动的多语知识共享研究”(2017JM7003)研究成果之一。

[通讯作者] 马晓悦(ORCID:0000-0003-4932-6450),博士,特聘研究员,研究方向:信息行为分析与交互设计,Email:xyama@mail@163.com。

[作者简介] 张思雅(ORCID:0000-0002-2914-6658),硕士研究生,研究方向:信息可视化管理,Email:syzhang@126.com; 赵捧未(ORCID:0000-0003-4026-6417),博士,教授,研究方向:信息存储与检索,信息资源管理,Email:pwzhao@mail.xidian.edu.cn。

## 1 引言

作为用户生成内容载体之一的社会化标签系统由于其开放性、交互性、社会性的特点成为用户信息交流和共享的主阵地<sup>[1]</sup>。社会化标签系统为用户提供了一个有效的信息资源描述方式,即用户不仅仅作为信息的接收者,更是信息生产过程的参与者,通过发表个人观点来组织和管理信息<sup>[2,3]</sup>。如著名的图片分享网站 Flickr,书签网站 Delicious 以及国内的电影图书分享网站豆瓣等<sup>[4]</sup>。近年来,经济全球化、社会多元化使得网络用户的国际化和网络信息的多语化趋势愈加明显,跨语言社会化标注系统受到广泛关注。然而,由于用户知识背景、标注习惯、信息素养的差异,以及这种“自下而上”的开放式信息组织方式所缺乏的顶层控制,标签系统中经常存在标签语义模糊、标签词汇冗杂等问题<sup>[5]</sup>。这些词汇问题对基于标签的信息组织和共享产生影响,增加用户寻找资源和发现主题的成本,成为影响社会化标签系统应用效果的重要因素之一<sup>[6]</sup>。特别地,在跨语言标注系统中,来自不同地区的用户使用相对独立的语言和思维标注信息资源,这进一步加剧了社会化标注中的词汇问题<sup>[7]</sup>。

推荐标签 (Suggested Tag, Recommended Tags) 被认为是解决社会化标签系统中词汇问题的有效方案之一,是由系统按照一定算法展示给用户的高质量标签<sup>[8]</sup>,按照其来源可以分为用户标注推荐标签和机器抽取推荐标签<sup>[9]</sup>。这两种不同来源的推荐标签都在用户标注过程中发挥着引导标注角度选择和主题发现的重要作用,不仅方便用户添加标签,而且提高了标签系统中的语义稳定性,改善了标签语义模糊、标签过于个性化及无意义标签泛滥等词汇问题<sup>[10]</sup>。

然而目前关于推荐标签的研究大多涉及推荐算法的优化<sup>[11]</sup>,鲜有学者关注推荐标签如何改变标注习惯和标签质量。研究推荐标签对用户标注行为的影响对于更好地完善系统功能具有重要意义。特别是在跨语言标注环境下,由于推荐标签表征语言和用户标注语言的多样性,推荐标签对用户标注行为的作用更加复杂。因此本研究通过受控实验的方法,探究跨语言环境下推荐标签及其语言对用户的

标注行为的影响,并关注用户在学习推荐标签进行标注过程中的心理感受和体验。

研究跨语言环境下用户对推荐标签的学习行为,一方面有助于挖掘推荐标签对用户行为的影响机理,从而反馈给推荐算法以优化推荐结果;另一方面可以了解用户在标注过程中的心理感受以提高跨语言标注系统中的用户标注积极性及用户体验。

## 2 国内外相关研究

社会化标签系统通过用户持续的标注行为,形成了“用户-标签-资源”三元关系网络<sup>[12]</sup>。标签作为用户与用户、用户与资源交互的桥梁,在社会化标注系统中发挥着主体作用<sup>[13]</sup>。研究标签如何产生,即用户作为信息组织者的标注行为,对提高社会化标注系统中的交互效率具有重要意义。已有研究证明,用户的标注经验、知识背景和认知特征以及资源类别、应用领域等都会导致用户的标注行为和标注结果存在差异。潘婵等发现用户对不同领域的资源进行标注时,所产生标签和文本关键词的语义相似度存在差异<sup>[14]</sup>。吴丹则关注不同应用环境下用户的标注行为,发现豆瓣图书标签和武汉图书馆用户标签在标签类型、标签形式、标签数量等方面存在显著差异<sup>[15]</sup>。标注过程中的用户行为是指用户具体的标注行为,如标注路径、标注策略等<sup>[16]</sup>。马费成等通过从豆瓣采集的真实数据构建标签共现网络,从而揭示用户在图书标注环境下的认知行为<sup>[17]</sup>。刘惠宇基于标签流行度及用户的知识背景构建出社会化标注动力学模型<sup>[18]</sup>。陆泉则关注图像标注中的用户行为,提出 10 种用户标注模式,以反映用户思考和确定标注结果的过程<sup>[19]</sup>。标注结果中体现的用户标注行为则从宏观角度分析用户的群体行为,区别于标注过程中的行为细节。大量研究从社会化标注系统的日志数据出发,挖掘用户的标注行为,结果发现标签出现的频率遵循幂律分布<sup>[20,21]</sup>。尽管在跨语言社会化标注环境中用户的知识背景、信息目标以及标注习惯等各不相同,他们却共享相同的行为,即学习之前用户添加的标签进行标注,从而提高社会化标注系统中标签的一致性<sup>[22]</sup>。

跨语言社会化标注行为是社会化标注行为的一

个分支,其研究内容和社会化标注行为具有重叠部分。但是,由于存在多种语言表征的标签,因此跨语言社会化标注行为研究更关注如何厘清不同语言标签之间的语义关系从而提高信息共享的效率。Jung 通过评估用户的外语水平和标注经验,提出基于多语标签的跨语言信息检索模型,摆脱了多语言字典的使用<sup>[23]</sup>。其次,研究多语言用户在标注时的语言选择行为和单语言用户的跨语言标注行为也是一个重要的研究方向。如国外学者 Steichen 等发现跨语言信息交互中用户大量使用多语标签,语言熟练度、使用目的以及主题领域等语境因素都对他们的语言选择和标签使用频率具有显著影响<sup>[24]</sup>。

推荐标签作为改善标签质量、降低用户标注难度的主要方法成为学者关注的一个研究内容。一方面大量学者研究基于推荐标签的信息推荐方法,包括基于标签概率、标签内容的推荐方法<sup>[25-27]</sup>。随着用户个性化信息需求的日益强烈,很多学者提出融入用户行为的标签推荐方法,如 Kowald 通过建立用户认知全过程的模型以优化标签推荐算法<sup>[28]</sup>。另一方面,学者还关注用户与推荐标签的交互行为,特别是标注过程中用户对推荐标签的学习与模仿。社会化标签系统中的基于推荐标签的标注学习行为是指在协同学习环境中用户对资源已有标签的学习模仿行为。其核心观点是个体通过观摩他人的行为而潜移默化地学习,最终形成新的行为或改变原有行为习惯的过程<sup>[29]</sup>。在社会化标注系统中,推荐标签包含用户标注对象选择、标注角度选择和标签表达等方面的有用信息,使得用户会选择跟从和学习推荐标签<sup>[30]</sup>,这种学习可以促进标签系统的标签质量提高,进而改善标签系统信息组织和共享的效率<sup>[31]</sup>。Fu 的研究发现社会化标签系统中用户从语义层面学习推荐标签从而提高标签的语义稳定性<sup>[32]</sup>; Floeck 进一步探究了标注过程中启动语义学习的边界条件<sup>[33]</sup>。

综上所述,之前学者从多个方面研究用户的标注行为,但是鲜有学者关注推荐标签对跨语言标注行为的影响机制。在跨语言环境下,由于推荐标签表征语言和用户标注语言的多样性,推荐标签对用户标注行为的作用更加复杂。因此,深入分析推荐

标签对跨语言标注行为的影响,关注用户在学习推荐标签进行标注时的心理感受是很有必要的。

### 3 跨语言环境下社会化标注学习的用户实验设计

基于以上分析,实验主要关注以下两个研究问题。首先,通过对比有推荐标签组和无推荐标签组的实验数据探究推荐标签对用户的跨语言标注行为和标签特征的影响。其次,通过对比中文推荐标签和英文推荐标签小组的实验数据研究不同表征语言的推荐标签对用户标注过程及用户体验的影响。因此,笔者招募母语为中文的大学生作为参与者,标注资源的表征语言和用户标注语言都为英文,设计了一个基于推荐标签的跨语言标注实验,包括英文推荐标签小组、中文推荐标签小组和无推荐标签小组。为保证被试者尽可能保持热情的标注兴趣,我们将标注实验控制在一个小时,并邀请4名实验人员进行模拟标注,发现在一个小时内用户最多标注不超过15篇文本。因此实验设计要求被试者在一个小时内对15篇外文资源进行跨语言标注。我们主要从三个层面对实验结果进行评估,旨在探究在跨语言信息标注中用户是如何学习推荐标签的,他们对不同语言表征的标签的学习模式和学习心理有无区别。在标签特征方面,本研究关注标签长度,即标签的单词数量。在标签语义方面,本研究关注标签来源及标签语义相似度指标。在标注行为方面,本研究关注用户添加的标签数量及标注速度。

#### 3.1 被试情况介绍

来自信息管理系的39名学生作为被试者,男女比例为1.6:1。他们被随机分为三个小组(每组13人):英文推荐标签小组、中文推荐标签小组和无推荐标签小组,分别记为A、B、C组。其中C组作为对照组。实验开始之前,通过先验问卷了解参与者的语言背景 and 知识背景(本研究的信息文本背景是旅游信息,因此需要了解他们对旅游信息的熟悉程度)。问卷采取李克特五级量表进行测量,最低分为0分,最高分为4分。

(1) 英语水平:75%的参与者都通过了全国大学生英语4级或6级考试。且当被问及“请对个人的英语水平进行评价”时,结果显示所有参与者的得

分都大于等于 1 分, 且三组用户的英语水平并无显著差异 (mean. A = 2.31, mean. B = 2.62, mean. C = 2.36, sig. = 0.506 > 0.05), 说明 3 组参与者的英语水平差异不会影响实验结果的可信性。

(2) 旅游知识背景: 三组用户对旅游领域的知识的了解程度处于中等水平, 且在国内旅游知识和国外旅游知识的了解程度方面都不存在显著差异 (国内: sig. = 0.961 > 0.05; 国外: sig. = 0.923 > 0.05), 说明了三组参与者的国内和国外旅游知识背景差异不会影响跨语言标注结果的可信性。

(3) 信息标注经验: 三组用户的信息标注经验平均处于基本了解的水平, 且不存在显著差异 (sig = 0.168 > 0.05), 保证了实验数据的可靠性。

综合而言, 本次实验中三组用户的外语水平、知识背景及信息标注经验等能力基本满足实验要求, 三组用户间不存在显著差异, 保证了实验结果的准确性。

### 3.2 实验材料

跨语言标注实验所使用的平台如图 1 所示。该平台允许用户在进行跨语言信息标注时查看系统提供的推荐标签 (如果有的话), 并为信息资源自由地添加标签。在标注过程中, 用户可以决定是否从推荐标签中选择标签进行标注, 或直接根据个人的信息目标和知识背景为其添加标签。

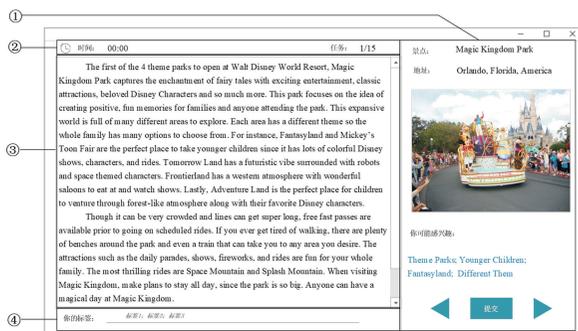


图 1 实验标注平台界面

①资源的关键信息: 景点名称、地址和推荐标签; ②标注时间记录; ③景点简介; ④标签添加栏

实验所用语料库来源于全球最大的旅游社区网站 TripAdvisor (<https://www.tripadvisor.com>, 英文

版), 包括景点信息简介以及推荐标签 (由之前用户添加的标签产生), 如图 2 所见。



图 2 三组实验材料展示

注: 从左到右分别为: 无推荐标签小组、中文推荐标签小组和英文推荐标签小组。

景点信息简介包括地址、特色、旅游攻略等内容。景点类型涵盖自然景观类、人文历史类, 娱乐公园等多种类型, 以满足不同实验者的偏好, 避免由于景点种类单一影响实验结果。为保证标注文档的信息丰富性和可读性, 笔者对收集到的景点简介的内容和长度进行调整, 在保证信息充分的情况下降低用户的阅读难度, 最终文档的长度控制在大概在 300 个单词左右。

推荐标签有两种语言表达形式: 英文推荐标签和中文推荐标签。其中英文推荐标签来源于该网站中由其他用户之前添加的标签集合, 平均为每篇标注文档提供四个英文推荐标签。中文推荐标签与英文推荐标签内容保持一致, 是通过机器翻译和人工翻译相结合的方式得到的。

### 3.3 实验方法及过程

实验用户的任务是在规定时间内, 在系统展示不同推荐标签的情况下为实验所提供的外文文本添加外文标签, 以实现资源的表征语言和用户标注语言都为外语的跨语言标注。具体实验过程为:

(1) 39 个参与者被随机分为三个小组, 记为 A、B、C 组 (详见被试情况介绍)。

(2) 研究人员向志愿者展示实验所用的平台,

告知参与者相关的操作以及标注任务。鼓励参与者积极提问,以保证他们在实验中能正常使用系统,从而避免由于实验者对实验平台不熟悉对实验造成的干扰。

(3) 参与者登陆系统并开始对系统中的外文文本按照编号依次进行标注。其中A组用户在标注过程中可以看到系统为其展示的英文推荐标签,B组用户可以看到系统提供的中文推荐标签,而C组用户在标注过程中仅仅获得系统提供的英文标注文本,没有任何推荐标签可供参考。

(4) 实验持续一个小时。当参与者完成所有资源的标注后点击“完成”按钮提交。或者当时间到达一个小时,系统直接退出用户的标注界面并保存数据。

#### 4 数据结果展示

作者对日志文件中的人员信息、标注时间、标签集合等数据进行清洗和分析。针对用户添加的标签字母大小写、单复数不统一等问题,使用文本处理方法对标签形式进行清洗和转换。

##### 4.1 标签的形式特征分析

本研究从标签长度这一变量入手分析标签的形式特征。已有研究发现,英语为母语的用户标注的英文标签长度不超过3个单词的标签占到总标签的98%以上<sup>[34]</sup>,这说明用户在标注时大多不会添加较长的标签,而是倾向于添加单词数量较少的短标签。而短标签在一定程度上较长标签更容易被重用,因此用户对短标签的流行性和接受度相对较高,有助于提高信息组织和共享的效率<sup>[35]</sup>。本实验在提供英文推荐标签时遵循这一发现,要求所有英文推荐标签的单词数量不超过三个。表1展示了不同长度标签的占比情况。分析发现A组单词数量不超过3的标签比例最大(A:93.47%;B:85.3%;C:86.98%)。这说明英文推荐标签小组的用户相较中文推荐标签组和无推荐标签组的用户在标注过程中整体添加的标签的长度更短。另外,我们发现三组中不超过三个单词的标签占比都小于上文提到的98%的结果。这说明用户使用外语进行标注和使用母语进行标注时标签的长度是有差异的。这是因为当用户在进行

跨语言标注时,其外语使用能力有限,导致他们在标注过程中难以使用很少的单词就能清晰地表达个人的观点,从而需要使用包含更多单词的标签。

表1 三组不同长度标签占比

小组 标签的 单词数量	A组	B组	C组
1	579(42.86%)	508(35.40%)	696(46.93%)
2	550(41.67%)	519(36.17%)	438(29.53%)
3	118(8.94%)	197(13.73%)	156(10.52%)
4	55(4.17%)	105(7.32%)	84(5.66%)
5~6	15(1.14%)	80(5.57%)	68(4.59%)
>6	3(0.22%)	26(1.81%)	41(2.77%)
总数	1320(100%)	1435(100%)	1483(100%)

其次,通过进一步对比三组标签的平均长度发现,英文推荐标签小组的用户产生的标签的平均单词数量为1.728( $SD=0.223$ ),而中文推荐标签小组的平均长度为2.190( $SD=0.258$ ),无推荐标签小组的平均单词数量为2.101( $SD=0.295$ )。这一变量通过了协同一致性检验( $sig.=0.837>0.05$ )。通过多重比较发现,英文推荐标签小组和中文推荐标签小组之间有显著差异( $sig.=0.023<0.05$ ),英文推荐标签小组和无推荐标签小组存在显著差异( $sig.=0.015<0.05$ ),而中文推荐标签小组和无推荐标签小组的标签长度之间没有显著差异( $0.357>0.05$ )。这个结果说明,英文推荐标签小组的用户相较中文推荐标签和无推荐标签小组的用户在标注过程中更倾向于使用更少单词的标签。特别地,在标注过程中能看到英文推荐标签的用户产生的标签的长度和英文推荐标签的长度具有相似性。这证明了标签长度作为一种标签的形式特征在标注过程中为用户提供了一定标准和参照,使得用户在标注过程中产生了基于标签形式的学习。相反地,中文推荐标签和无推荐标签由于不能为用户提供形式上的指导,导致用户难以选择合适的词作为标签,只能通过包含更多单词的标签以满足个人的信息目标。因此,为用户提供与其标注语言相同的推荐标签可以为用户提供标注规范,提高标注质量和用户体验。

##### 4.2 标签的语义特征分析

分析用户添加的标签和推荐标签间的语义关系

为研究用户基于标签语义的学习行为提供有力的证据,能够有效地反映和揭示用户的学习行为。潜在语义分析(Latent Semantic Analysis)是通过分析大量文本来提取和表示单词和段落语义相似性的一种代数模型和统计技术。本研究采用潜在语义分析方法(Latent Semantic Analysis, LSA)测量参与者添加的标签集合和推荐标签集合间的语义相似度,不仅因为该方法的实现比较容易,还因为之前的研究为使用LSA反映单词语义相似性提供了很好的验证<sup>[36]</sup>。我们通过网站 <http://lsa.colorado.edu> 进行LSA计算,最后两个标签的LSA值通过其行向量的cos值表示<sup>[37]</sup>。由于B组的推荐标签是由中文表征的,因此不能直接与用户添加的英文标签进行语义计算。因此本研究对中文推荐标签采取先翻译再计算语义相似度的方法,即使用A组的英文推荐标签代替中文推荐标签以计算B组标签间的语义相似性。表2分别展示了英文推荐标签小组和中文推荐标签小组每一篇文档的参与者产生的标签集合与推荐标签间的语义关系。通过方差分析我们发现,英文推荐标签小组的平均语义相似性显著高于中文推荐标签小组( $\text{sig.} = 0.044 < 0.05$ )。通过比较两组的平均语义相似性发现,英文推荐标签小组的平均LSA值大于中文推荐标签小组(A:  $\text{mean} = 22.657$ ; B:  $\text{mean} = 17.463$ )。这个结果说明,在英文推荐标签环境下用户添加的标签与推荐标签之间的语义关系更近,而在中文推荐标签条件下标签与推荐标签的语义关系较远。

为了进一步探究推荐标签的语言对用户语义学习行为影响的深层原因,笔者进一步分析了各小组标签产生的来源,以反映用户在基于语义添加标签时的具体过程。本研究将标签的来源大致划分为两类,第一类为用户直接从外文文本中挑选与推荐标签语义相近的关键词作为标签,第二类为用户个性化创造的标签(开放式标签)。这两种来源分别对应两种不同的学习行为。前者语义映射的范围较小、层次较浅,因此与推荐标签及标注文本的语义关系较近。后者是通过启动个人的理解过程,结合资源信息和个人知识背景,开放式地从多个角度添加

满足个人信息目标的标签,因此使得语义映射范围较广、层次较深,与推荐标签及标注文本的语义关系较远。

表2 A组(英文推荐标签小组)和B组(中文推荐标签小组)标签的LSA分数

文本	A组	B组
文本 1	16.367	15.783
文本 2	28.931	23.572
文本 3	26.603	19.864
文本 4	32.759	25.634
文本 5	28.307	24.695
文本 6	24.751	16.279
文本 7	23.632	15.766
文本 8	30.325	23.831
文本 9	22.758	17.538
文本 10	27.143	21.839
文本 11	21.658	18.713
文本 12	24.149	16.261
文本 13	12.413	8.793
文本 14	10.185	6.951
文本 15	9.871	6.430

采用文本分析的方法分别计算三组用户添加的标签的来源,即通过计算用户产生的标签集合中与标注文本和推荐标签不重合的单词数量来计算开放式标签的占比,如图3所示。通过方差分析我们发现,B组由用户自由创造的标签的占比显著高于A组( $\text{sig.} = 0.031 < 0.05$ ),而A组和C组( $\text{sig.} = 0.462 > 0.05$ ),B组和C组之间却不存在显著差异( $\text{sig.} = 0.133 > 0.05$ )。即中文推荐标签组的用户倾向于添加更多自由创造的开放式标签,而英文推荐标签组的用户则倾向于从文档中选择与推荐标签语义相近的关键词作为标签(A:  $\text{means} = 22.9\%$ ,  $\text{SD} = 0.071$ ; B:  $\text{means} = 25.9\%$ ,  $\text{SD} = 0.094$ ; C:  $\text{means} = 24.53\%$ ,  $\text{SD} = 0.048$ )。

根据以上分析结果,得到两点启示:①在进行跨语言标注时,用户主要是基于推荐标签从文本中选择关键词进行标注,而在跨语言标注时自由创造个性化标签的行为并不像在对熟悉的语言进行标注时结合个人的信息目标和背景知识进行开放式标注以满足个人信息需求那样普遍。②推荐标签的语言对用户进行跨语言标注时的学习行为具有显著影响。

具体来说,能够看到中文推荐标签的用户在跨语言标注时相比能看到英文推荐标签的用户来说,个性化创造标签的行为更加普遍。而能看到英文推荐标签的用户更加偏向于从文本中挑选与推荐标签语义相近的关键词作为标签。这一结果与标签间的语义分析的结果正好一致。因为在跨语言标注环境下,由于外语熟练度较低,用户在理解外文文本和外文推荐标签时存在语言障碍,因此他们进行启动理解过程和主观创造标签的行为被抑制,从而导致在标注过程中更多地直接选择和文档主题相关的关键词作为标签,这种直接地产生标签的方式使得英文推荐标签小组中用户产生的标签集合和推荐标签间的语义相似性更大。而中文推荐标签小组地推荐标签由于以用户的母语表征,在一定程度上相比英文推荐标签可以帮助用户理解文本信息,因此用户会自发地启动语义解释过程,结合个人的知识背景和信息目标创造更加能满足个性化信息需求的标签,从而使得中文推荐标签小组中由用户产生的标签集合与中文推荐标签间的语义相似性更小。

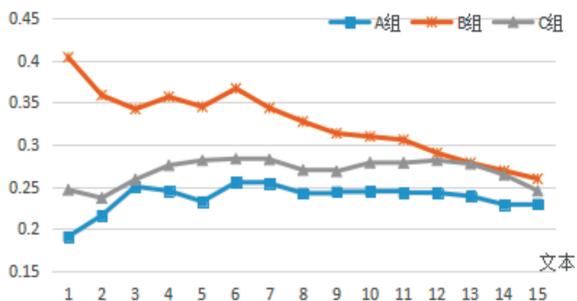


图3 三组开放式标签占比情况

#### 4.3 标签数量与标注时间分析

分析各组标签集合的标签数量和用户标注时间可以从宏观的角度分析用户的群体行为。结果显示标注实验中A组用户共添加了1320个标签,B组用户共添加了1,435个标签,C组用户共添加了1,483个标签。通过方差分析发现,A组和C组的标签数量间有显著差异( $\text{sig.} = 0.014 < 0.05$ ),B组和C组也具有显著差异( $\text{sig.} = 0.002 < 0.05$ ),A组和B组的标签数量都显著少于C组(A:  $\text{mean} = 8.557, \text{SD} =$

$0.917$ ; B:  $\text{mean} = 8.991, \text{SD} = 0.820$ ; C:  $\text{mean} = 10.463, \text{SD} = 1.46$ )。而A组和B组的标签数量间不存在显著差异( $\text{sig.} = 0.534 > 0.05$ )。这一结果与之前关于单语言环境下的研究结果刚好相反<sup>[38]</sup>。之前研究已经证明当为用户提供推荐标签时,用户在选择标注角度、单词时会更加容易,因此用户会添加更多标签。然而本研究发现当为用户提供外文推荐标签和中文推荐标签时,用户所添加的标签数量都显著少于没有推荐标签的情况。

本研究的标注时间是通过标注平台记录的标注每一篇文档的时间比用户所添加的标签数量得到的。标注时间反映了平均添加一个标签所用的时间。标注时间计算公式为:

$$TT = \frac{\sum_{i=1}^{13} T/13}{\sum_{i=1}^{13} TN/13} \quad (1)$$

TT (Tagging Time) 表示标注时间, T (Time) 表示每个用户标注每篇文本所花费的时间, TN (Tags Number) 代表每个用户为每篇文本所添加的标签的数量(13代表每个小组中参与者的数量)。

由于标注实验具有时间限制,因此并不能保证所有用户在实验结束之前完成所有文档的标注任务。这导致编号靠后的文档并不能被所有用户标注,为了保证实验数据的有效性,本研究选取前8个文档的数据进行分析(因为观察数据发现所有用户都完成了前8篇文档的标注任务)。结果发现A组用户平均标注一个标签的时间是38秒,B组为32秒,C组为34秒。通过方差分析发现A组和B组的标注时间存在显著差异( $\text{sig.} = 0.018 < 0.05$ ),而A组和C组,B组和C组间不存在显著差异(A and C:  $\text{sig.} = 0.173 > 0.05$ ; B and C:  $\text{sig.} = 0.499 > 0.05$ )。这个结果说明:相对于没有推荐标签的小组,可以看到英文推荐标签的用户在标注时的速度更慢,而能看到中文推荐标签的用户标注的速度更快。

## 5 结果讨论

### 5.1 基于推荐标签的跨语言标注学习模式

学习模式是指在跨语言标注中用户所使用的学习方法和学习行为的组合。根据实验数据分析结果

发现, 基于推荐标签的用户学习方法主要包含两种, 即基于标签形式的学习和基于标签语义的学习。前者仅仅体现在标注语言和推荐标签的表征语言一致的情况下。推荐标签的可视化特征或作为一种符号指导用户的学习过程, 在该学习方法中标签的语义不发挥作用。后者与标签的表征形式无关, 用户仅仅关注标签的语义, 根据语言熟练度对语义理解程度的不同而产生两种不同的学习行为: 长语义映射学习行为和短语义映射学习行为(如图4所示)。长语义映射行为是指, 当用户对推荐标签的语言比较熟练时, 用户可以快速地抽取和归纳标注资源的主题, 获取推荐标签的标注视角, 从而启动融入个人知识背景和信息目标的理解过程, 开放性地添加更多个性化标签。短语义映射学习行为是指当用户对推荐标签的表征语言不够熟悉时, 用户难以快速获取全面有效的关于推荐标签的内容, 从而缺乏充分发挥个人理解的启动条件。因此根据“最省力原则”<sup>[39]</sup>, 用户倾向于使用类比、匹配等方法直接从文档中选择与推荐标签语义相近的关键词作为标签, 而很少基于个人理解个性化地创造新的标签。基于标签来源的数据分析结果说明 B 组相比 A 组具有更多开放式标签, 即中文推荐标签小组的参与者创造了更多个性化标签, 而英文推荐标签小组的参与者倾向于从标注文本中选择关键词进行标注。潜在语义分析结果也证明了 A 组相比 B 组其用户标注的标签和推荐标签间具有更强的语义结构, 这一结果也验证了中文推荐标签相比英文推荐标签更能刺激用户产生开放式的语义联想。因此, 基于推荐标签的跨语言标注学习模式共包括三种: 基于标签形式的学习模式、基于标签语义的长语义映射学习模

式以及基于标签语义的短语义映射学习模式。

## 5.2 基于推荐标签的跨语言标注学习心理

本研究进一步探究了不同推荐标签在用户标注学习过程中对用户心理的影响。实验结果表明, 在用户的跨语言标注过程中, 能看见推荐标签的用户添加的标签数量相对较少, 且在外文推荐标签条件下, 用户的跨语言标注时间也较长, 语言不熟悉的推荐标签为用户标注带来了更大代价, 从而导致标注速度较慢。而中文推荐标签因为用户对其表征语言比较熟悉, 在帮助用户语义理解方面体现出一定的优势, 降低了用户的标注难度, 提高了用户的标注速度和标注体验。这说明用户在学习推荐标签时, 不仅仅获得其带来的线索和指导以帮助自己理解文本内容、总结文本主题以及选择合适的标签。特别地, 在本研究中我们发现推荐标签还为用户的跨语言标注学习带来一定的代价, 这是与之前单语言标注环境下研究发现的的不同之处。通过进一步分析发现, 这种代价伴随着推荐标签为用户提供指导而发生。借鉴心理学领域的锚定效应(Anchoring Effect)可以更具体地解释这一发现。

锚定效应认为在用户进行决策时, 外界已有的重要信息会被用户无意识地认为是一个标准, 这个标准相当于“锚”一样制约着用户的决策。锚定效应认为“锚”只要受到人们的注意, 那么无论其数据是否具有实际参考效用, 都会对用户做出决定具有一定影响<sup>[40]</sup>。用户在做决定时会不自觉地依赖这个“锚”做出估测。在跨语言标注中, 推荐标签则刚好扮演了“锚”的角色, 使得用户在标注过程中产生了对“锚”的学习行为。与此同时推荐标签也为用户带来了一定的限制, 导致用户在跨语言标注学习时更加严苛, 以尽可能添加和“锚”一致的信息。之前的研究已经发现, 当用户产生更多与“锚”一致的知识时, 甚至体现出更强的锚定效应时, 他们会更加感受到一种悲伤的负面情绪<sup>[41]</sup>。因此这种更加精细的对推荐标签的学习行为使得用户付出一定的代价, 如心情低落或心理压力增大, 从而导致在实验中他们的标注速度减慢、标注数量减少。

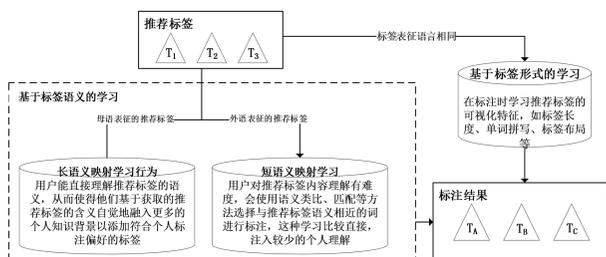


图4 基于标签的跨语言标注学习模式

## 6 总结

本研究关注基于添加标签的跨语言信息标注过

程中的用户学习行为,并设计了一个包含英文推荐标签小组、中文推荐标签小组以及无推荐标签小组的跨语言标注实验,了解不同小组的用户在整个标注过程中的行为特征,进而寻找推荐标签特征与标注行为的关系。通过基于推荐标签的学习模式和学习心理两个方面分析用户行为。前者关注跨语言标注过程中的具体行为与方法,后者关注在学习过程中用户的心理感受。一方面,研究结果发现跨语言环境下存在三种具体的学习模式。首先,在基于推荐标签形式的学习模式中,当推荐标签和用户标注的表征语言相同时,推荐标签的外在可视化的特征会对用户标注发挥一定的指导作用。这种基于标签形式的学习模式的发现为之后的研究人员提供了一种提高标签质量的新视角。具体来说,标签与文本内容和信息目标的匹配度不是判断推荐标签质量的唯一标准,标签的可视化特征也是对推荐标签质量进行评价的一个标准,如标签布局、标签长度、单词拼写等形式特征。其次,实验分析发现的另外两个标注学习模式与基于标签形式的标注模式有着根本性差异。笔者提出在基于推荐标签语义的跨语言标注学习模式中,由于用户对推荐标签的表征语言熟练度不同存在着长语义映射学习模式和短语义映射学习模式。长语义映射学习是指基于推荐标签融入个人知识背景理解的更加开放的学习过程,而短语义映射模式是指基于推荐标签的词汇匹配和类比的直接学习过程。另一方面,本研究分析了不同推荐标签环境下用户学习的心理感受。最终发现用户对标签的学习符合锚定效应。更重要的是,跨语言标注情境下用户在学习推荐标签的同时,他们也会付出一定的代价,这种代价体现在标注速度和标注数量的减少,以及在标注模仿时的心理压力感和焦虑感。这一发现在肯定之前研究关于推荐标签优点的研究发现的同时,创新性的发现推荐标签使得用户在标注过程中也需付出一定代价。

以上这些用户跨语言标注学习行为的分析将有助于指导跨语言标注系统的设计与实践,提高用户满意度和信息共享效率。此外,笔者认为跨语言标注情境下学习行为研究还有进一步深入的空间。探

究更多标签的形式特征对用户标注学习的影响是未来研究的一个方向。在今后的研究中,笔者将扩大研究范围,选择背景更多元化的用户参与实验,更加深入探究跨语言标注情境中的用户学习行为。

#### 作者贡献说明

张思雅:设计实验方案,实施实验,采集、清洗和分析数据,论文撰写;

马晓悦:提出研究思路和整体研究框架,设计研究方案,论文最终版本修订;

赵捧未:论文最终版本修订。

#### 支撑数据

支撑数据由作者自存储,Email:syzhang@126.com。

- 1 张思雅.Pre-questionnaire.xlsx.先验问卷数据。
- 2 张思雅.Tags.xlsx.用户标签数据。
- 3 张思雅.Tagging time.xlsx.用户标注时间数据。

#### 参考文献

- 1 李旭晖,李媛媛,马费成.我国图情领域社会化标签研究主要问题分析[J].图书情报工作,2018,62(16):120-131.
- 2 邵杨芳,陈新国.社会化标注系统中的知识协同效率影响因素模型研究[J].情报学报,2017,36(4):361-369.
- 3 李蕾,王冕,章成志.区分标签类型的社会化标签质量测评研究[J].图书情报工作,2013,57(23):11-16,9.
- 4 吴丹,许小梅.图书馆与图书分享网站的用户标注行为比较研究[J].图书情报知识,2013(1):85-93.
- 5 熊回香,杨雪萍.社会化标注系统中的个性化信息推荐研究[J].情报学报,2016,35(5):549-560.
- 6 Mahatab U, Jabeen F. Spam Tags Detection and Protection Using Tags Relationship Based Anti-Spam Approach[C]//2018 IEEE 21st International Multi-Topic Conference (INMIC). Karachi:IEEE,2018:1-7.
- 7 徐少同.网络信息自组织视角下的 Folksonomy 优化[J].图书情报工作,2009,53(10):102-105,120.
- 8 贾君枝,孙智超,邵杨芳.基于受控词表的医学资源社会化标签推荐研究[J].情报学报,2013,32(12):1326-1332.
- 9 顾晓雪.多语言社会化标签聚类及可视化研究[D].南京:南京理工大学,2015.
- 10 Kopeinik S, Lex E, Seitlinger P, et al. Supporting Collaborative Learning with Tag Recommendations: A Real-world Study in an

- Inquiry-based Classroom Project[C]//Proceedings of the Seventh International Learning Analytics & Knowledge Conference. New York: ACM, 2017: 409-418.
- 11 Seitlinger P, Kowald D, Trattner C, et al. Recommending Tags with a Model of Human Categorization[C]//Proceedings of the 22nd ACM international conference on Information & Knowledge Management. San Francisco: ACM, 2013: 2381-2386.
- 12 潘旭伟,傅青苗. 基于超网络的社会化标注行为[J]. 系统工程, 2015(3):78-83.
- 13 Gursoy A, Wickett K, Feinberg M. Understanding Tag Functions in a Moderated, User-generated Metadata Ecosystem[J]. Journal of Documentation, 2018, 74(3):490-508.
- 14 潘婵,冯利飞,丁婉莹,贺芳. 基于标签-关键词的用户行为分析[J]. 情报杂志, 2010, 29(3): 139-142.
- 15 吴丹,许小梅. 图书馆与图书分享网站的用户标注行为比较研究[J]. 图书情报知识, 2013(1): 85-93.
- 16 Tsai L C, Hwang S L, Tang K H. Analysis of Keyword-based Tagging Behaviors of Experts and Novices[J]. Online Information Review, 2011, 35(2): 272-290.
- 17 马费成,张斌. 图书标注环境下用户的认知特征[J]. 中国图书馆学报, 2014, 40(1): 4-14.
- 18 刘惠宇. 基于标签流行度的社会化标注模型研究[D]. 大连: 大连理工大学, 2014.
- 19 陆泉,韩雪,陈静. 图像标注中的用户标注模式与心理研究[J]. 情报学报, 2015, 34(5): 125-129.
- 20 卢超,章成志. 中英译本图书社会化标签的比较研究[J]. 图书情报工作, 2013, 57(23): 17-23.
- 21 吴振宇,胡军,李德毅. 社会标注系统幂律特性分析[J]. 复杂系统与复杂性科学, 2014, 11(2): 5-16.
- 22 王向前,李慧宗. 基于资源内容聚类社会化标签聚类方法[J]. 情报杂志, 2016, 35(11): 141-145, 150.
- 23 Jung J J. Discovering Community of Lingual Practice for Matching Multilingual Tags from Folksonomies[J]. The Computer Journal, 2011, 55(3): 337-346.
- 24 Steichen B, Ghorab M R, O' Connor A, et al. Towards Personalized Multilingual Information Access-exploring the Browsing and Search Behavior of Multilingual Users[C]//International Conference on User Modeling, Adaptation, and Personalization. Cham: Springer, 2014: 435-446.
- 25 魏建良,朱庆华. 社会化标注理论研究综述[J]. 中国图书馆学报, 2009, 35(6): 88-96.
- 26 张斌,张引,高克宁,等. 融合关系与内容分析的社会标签推荐[J]. 软件学报, 2012, 23(3): 476-488.
- 27 张亮. 基于 LDA 主题模型的标签推荐方法研究[J]. 现代情报, 2016, 36(2): 53-56.
- 28 Kowald D. Modeling Cognitive Processes in Social Tagging to Improve Tag Recommendations[C]//Proceedings of the 24th International Conference on World Wide Web. Florence: ACM, 2015: 505-509.
- 29 冯娇,姚志. 基于社会学习理论的在线评论信息对购买决策的影响研究[J]. 中国管理科学, 2016, 24(9): 106-114.
- 30 林鑫,梁宇. 用户社会化标注中非理性行为的表现及原因分析[J]. 数字图书馆论坛, 2016(12): 48-53.
- 31 Nwana A O, Chen T. Towards Understanding User Preferences from User Tagging Behavior for Personalization[C]//2015 IEEE International Symposium on Multimedia (ISM). Miami: IEEE, 2015: 178-183.
- 32, 36 Fu W T, Kannampallil T G, Kang R, et al. Semantic Imitation in Social Tagging[J]. ACM Transactions on Computer-Human Interaction, 2010, 17(3): 1-37.
- 33 Floeck F, Putzke J, Steinfels S, et al. Imitation and Quality of Tags in Social Bookmarking Systems-Collective Intelligence Leading to Folksonomies[M]//On Collective Intelligence. Berlin: Springer, 2010: 75-91.
- 34 Guyot A. Understanding Booksonomies: How and Why are Book Taggers Tagging? [D]. Aberystwyth: University of Wales, 2013.
- 35 张颖怡,章成志,池雪花,李蕾. 科研用户博文关键词标注行为差异研究——以科学网博客为例[J]. 现代图书情报技术, 2015(10): 13-21.
- 37 Blackmon M H, Kitajima M, Polson P G. Tool for Accurately Predicting Website Navigation Problems, Nonproblems, Problem Severity, and Effectiveness of Repairs[C]//Proceedings of the 2005 Conference on Human Factors in Computing Systems. Portland, USA: ACM, 2005, 31-40.
- 38 Gueye M, Abdessalem T, Naacke H. A Parameter-free Algorithm for an Optimized Tag Recommendation List Size[C]//Proceedings of the 8th ACM Conference on Recommender Systems. Foster, USA, ACM, 2014: 233-240.
- 39 任忠忠,胡德华. 最省力原则视角下的移动搜索行为研究[J]. 情报科学, 2016, 34(1): 96-103.
- 40 唐卫海,徐晓惠,王敏,等. 锚定效应的产生前提及作用机制[J]. 心理科学, 2014, 37(5): 1060-1063
- 41 Strack F, Bahk Š, Mussweiler T. Anchoring: Accessibility as a Cause of Judgmental Assimilation [J]. Current Opinion in Psychology, 2016, 12: 67-70.

(收稿日期: 2019-12-30)

## 跨语言标注情境中的学习行为——基于推荐标签的实验分析

Study on the Learning Behavior in the Context of Cross-language Tagging: An Experimental Analysis Based on Recommended Tags

张思雅 马晓悦 赵捧未

## 专家评审意见

本文对用户标注行为进行了对照实验分析,并尝试对学习行为进行总结。这一研究很有意义,但是在逻辑思路和写作上仍有需要改进的地方,尤其是实验的流程和实验的设计。

1. 第二部分:研究综述部分对文献有总结,但是评论需要加强,与本文研究密切相关的研究还不够,建议补充。

2. 第三部分:实验设计过程,3.1 实验人员,建议改为“被试情况介绍”。另外,“说明实验者的英语水平不会影响实验标注结果。”,研究人员对被试的英语水平的测试,发现3组被试之间不存在显著差异。这个控制不能说明英语水平不影响标注结果,而是标注结果的差异应该不是由于3组被试的英语水平的差异导致的。对旅游知识背景的表述也有同样的问题。

3. 第四部分:实验结果的介绍部分,逻辑不够清晰。数据的收集与分析不要写在一起,数据收集的时候考虑到自己的研究问题,提前说明要收集哪些用户标注行为,以及这些标注行为如何测量,如何比较。然后再根据研究问题,逐个介绍研究结论。另外,下面几个细节没有介绍:①LSA值如何计算的?没有给公式;②标签产生的来源的划分有文字的描述,但是没有给出具体的公式;③实验中每组被试标注的文本数量没有介绍;④方差分析具体是什么分析?应该有具体的说明,对结果的展示和比较建议用表。

4. 第五部分对用户学习模式的分析挺有意思的,如果可能的话建议结合用户的行为一起放在研究结论部分进行分析说明,然后再尝试分析和总结用户的学习模式。

(上接第42页)

## 专家评审意见

本文采用实验方法探索了用户在数字图书馆情境下搜索停止规则的使用情况及其可能的影响因素。文章整体内容完整,结构较为清晰,但在专业术语的使用和部分内容的组织安排上存在一定的问题。

1. 文章中存在“认知停止规则”和“搜索停止规则”反复交替使用的情况,建议文章在这一术语的使用上做到前后统一。

2. 文章在相关研究现状部分提到“关于信息搜索停止原因的研究,也称为搜索停止规则研究”,但在该领域研究中,搜索停止规则是由用户的停止行为而产生,而搜索停止行为的原因则涉及环境等外在因素和用户的内在心理等因素。那么,“搜索停止原因”和“搜索停止规则”是否为等同的两个概念?文中多处存在这两种说法混用的现象。但文章的核心内容则是对搜索停止规则的使用情况进行探索,并未探究搜索停止行为的原因。建议明确这两者之间的区别,并选择更为合适的说法。

3. 文章的研究内容为“搜索停止规则的使用情况”和“搜索停止规则使用的影响因素”,而文章的题目中“停止规则的使用行为”则较为模糊,请作者考虑现有题目是否可以很好地对文章的研究内容进行概括?

4. 文章3.4 任务设计与数据收集部分,建议给出这五类任务的设计原则或依据,否则单从现有的表述上来看,会让人产生疑问,例如为什么该类任务的设计是这样的?某两类任务的设计上有什么区别?另外,在该节提到了用问卷收集用户的“检索自我效能”,但“检索自我效能”这一变量在后续数据分析部分从来没有出现过,那用问卷收集用户“检索自我效能”的目的何在?

5. 文章4.1 认知停止规则使用的基本情况中“当被试对规则使用评分大于3时,认为该认知停止规则为用户实际使用”,是否有一定的依据?